ChatGPT 与生成式人工智能的兴起：对学术诚信的挑战？

杨波

（上海交通大学电子工程系，上海 200240）

摘要：随着开放人工智能公司（OpenAI）开发的 ChatGPT 的问世，公众对生成式人工智能（Gen AI）系统及其对学术诚信的潜在影响表现出了极大的关注。本文旨在综述当前关于 ChatGPT 与学术诚信的讨论，并得出结论：ChatGPT 及其他类似生成式人工智能系统的应用可能会对学术诚信造成损害。鉴于此，为了有效降低对学术诚信的威胁，需要机构层面的更大努力以及多方利益相关者的协作。

关键词：ChatGPT;大型语言模型;开放人工智能;学术诚信;生成式人工智能

The Rise of ChatGPT and Generative Artificial Intelligence: A Challenge to Academic Integrity?

Yang Bo

(Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240)

Abstract: With the introduction of ChatGPT developed by OpenAI, the public has shown great concern about generative artificial intelligence (Gen AI) systems and their potential impact on academic integrity. This paper aims to review the current discussions on ChatGPT and academic integrity, and concludes that the application of ChatGPT and other similar generative AI systems may damage academic integrity. Given this, greater efforts at the institutional level and collaboration among multiple stakeholders are needed to effectively mitigate the threats to academic integrity.

Keywords: ChatGPT; Large Language Models; OpenAI; Academic Integrity; Generative Artificial Intelligence;

1.简介

开放人工智能公司的 ChatGPT 的出现引起人们对生成式人工智能（Gen AI）系统及其对学术诚信的可能影响的高度关注。生成式人工智能系统的设计目的是从训练的数据中生成内容（如文本、图像、音频、仿真、视频和代码）。尽管 ChatGPT 既不是有史以来开发的第一个生成式人工智能系统，也不是开放人工智能公司的第一个，但它代表了生成式人工智能技术的突破[1]。在许多学术领域，人们对学术诚信提出了担忧。现有许多不同内容的人工智能文本/内容生成器，问题是为什么 ChatGPT 会引起高度关注呢？

为了获得更清晰的视角，首先科普什么是 ChatGPT。ChatGPT 是一个大型语言模型（LLM），它使用深度学习来生成自然语言回答用户输入的问题。它于 2022 年 11 月 30 日发布，是开放人工智能公司对能够进行“智能”对话的大型语言模型的最新迭代。它是总部位于加利福尼亚州的公司的生成式预训练转换器（GPT）模型的一部分。在此之前，已经有 GPT-1 于 2018 年发布、GPT-2 和 GPT-3。2021 年，开放人工智能公司发布了 DALL.E2，这是一个从文本中生成图像的第二代人工智能系统。然而，ChatGPT 在许多方面与以前的模型不同。最重要的

是，它与 GPT-3 不同，旨在执行广泛的自然语言处理（NLP），如语言翻译、文本总结和问题回答，生成创造性写作（如诗歌或小说），生成高质量的长或短文（如博客文章）。另一方面，ChatGPT 是基于 GPT-3 语言模型构建的，具有独特的使用场景（如对话/会话的回应，复杂问题、概念或主题的解释，新代码的生成或现有错误代码的修复）。总的来说，ChatGPT 比 GPT-3 有更多的应用场景，与许多其他技术一样，它具有逻辑延展性，这意味着它可以针对各种语言任务进行微调。ChatGPT 的能力被支持者誉为相当出彩，或被描述为先进、高效且仍在学习。此外，与许多人工智能驱动的内容生成器不同，它对所有用户都是免费的。ChatGPT 的固有能力已在报道中得到证明，它已成功通过法学院考试[2]和工商管理硕士（MBA）考试[3]。

然而，批评者指出，作为一种大型语言模型，ChatGPT 不是特别创新和具有革命性，因为过去也开发过类似的系统。其他人观察到，尽管该系统的文本流畅而有说服力，但它仍然缺乏真正理解人类语言和对话复杂性的能力。为了公平地对待开发人员，系统的一些局限已经向用户明确，它偶尔会产生不正确的信息，产生有害的指令或有偏见的内容，并且受限于它训练所接受的数据而知识有限。在 ChatGPT 引起的诸多问题中，这篇评论只探讨了它是否破坏了学术诚信。本文还就学术界如何积极应对 ChatGPT 和生成式人工智能系统带来的挑战提出了建议。

2. 对学术诚信的挑战？

在学术界，ChatGPT 已被证明能够有效地回应本科生和研究生常见的询问，包括那些需要编程技能的问题。然而，存在一种普遍的忧虑，即学生和研究人员可能会倾向于将他们的写作任务委托给 ChatGPT 来完成。根据一些早期大学级别的论文问题的回答，大学老师应该考虑论文作为一种评估方式的前景。

国际学术诚信中心（2021）在其第三版《学术诚信的基本价值观》[4]中明确指出，学术诚信要求个体坚守六大核心价值：诚实性、信任度、公正性、尊重他人、责任感以及勇于担当。因此，任何借助自动化文本生成工具如 ChatGPT 来产出论文或其他文本，并将其伪造成个人原创成果的行为，均被视为违反了学术诚信的根本准则。ChatGPT 引发了在高等教育中与证据充分的商业“合同欺诈”类似的担忧[5]。唯一的区别是 ChatGPT 是免费的，所有用户都可以轻松访问。它还为用户提供了互动的机会。用户可以调整他们的查询内容，以了解生成内容的差别。这意味着有可能生成不同的文本或文章，用户可以从中挑选出最好的。使用 ChatGPT 进行学术论文或作业是否会构成“侵犯知识产权”道德意义上的剽窃也是至关重要的。当 ChatGPT 被当做原创作品时，侵犯了谁的知识产权？这一行为损害了谁？虽然在没有适当引用的情况下使用 ChatGPT 违背了学术诚信的基本价值观，但整个剽窃的辩论要复杂一些。

围绕 ChatGPT 及其现实能力的炒作可能会让学术界人士感到震惊或兴奋。学术界关注的不仅仅是技术的好坏。ChatGPT 是一种颠覆性技术的代名词。这项创新有望彻底改变学术、科学和教育领域的基本理论和认知框架。这预示着知识界将不得不重新审视知识的本质及其获取方式。随之而来的挑战是，如何确保这项技术被有效地、安全地以及负责任地整合和运用。本评论的重点并不是讨论 ChatGPT 是一种道德中立的技术，还是标准道德秩序中的一部

分。这并不意味着 ChatGPT 不会引起学术诚信问题之外的其他伦理问题，或者这些担忧无关紧要。文献中已经发现了许多围绕大型语言模型的伦理问题[6]。关于建造 ChatGPT 的人力成本引起了人们的极大关注。尽管本文并不专注于此，但根据目前对这项新兴技术当前的了解，其潜在的应用可能会对学术诚信造成严重威胁。这引出了一个关键问题：学术界应如何应对这一挑战？

3.学术界能做什么？

在学术领域，研究者需关注众多事项，这不仅包括探索 ChatGPT 等大型语言模型所带来的潜在机遇，也需考虑这些模型可能带来的挑战。研究者必须掌握如何高效地利用这些技术提供的优势，同时还需设法减轻这些技术对学术诚信可能产生的负面影响。

3.1.考虑 ChatGPT 在学术界的机遇和挑战

学术界需要认真对待 ChatGPT。学术界是一个致力于推动研究、教育和学术探索的综合性环境。它涵盖了高等教育机构、研究中心、学术期刊出版机构以及资助机构。像 ChatGPT 这样的生成式人工智能技术，以及其他类似的智能系统，正在引领一场重大的技术革新。学术界必须适应并融入这一变革之中。简单地禁止、排斥或否认这些技术的存在，并不是一个可行的长期策略。这是一项为教学、研究和创新提供机会的技术。利用 ChatGPT 作为学术研究的辅助手段，不仅能够显著提高工作效率，还能极大地节省时间。该工具在教育领域的应用范围广泛，包括但不限于课程大纲的制定、学习任务的设计与布置、学术写作的辅助以及创意和观点的激发。ChatGPT 的介入，对于教师而言，能够提升教学质量；对于学生来说，则能够丰富学习体验。此外，它在学术研究方面的潜力同样不容小觑，有助于改进研究方法和流程。例如，它可以为许多类型的研究快速轻松地生成数据。此外，该工具还能作为分析手段，辅助撰写科学研究报告。

在学术界，确保负责任地运用人工智能工具，如 ChatGPT，是一个日益重要的议题。特别是，我们必须警惕其潜在的滥用行为，这些行为可能会对学术诚信造成严重的威胁。

首先，在学术领域，恰当地引用他人的研究成果是遵守学术诚信的关键。尽管如此，一些学术工具的正确使用方式尚未被学术机构和出版商的诚信政策所涵盖。同时，学术界的众多成员，如研究者、教育者和学生，对于如何有效地利用这些工具，以及如何以负责任的态度进行操作，仍然知之甚少。因此，提供教育和培训变得至关重要。

其次，尚未建立一种统一和负责任的方式来引用 ChatGPT。许多研究论文将 ChatGPT 列为作者[7]。然而，《自然》和《科学》不允许大语言模型作为杂志认可的作者。目前缺乏对用户如何引用 ChatGPT 的指导引发了很多担忧。

第三，在学术领域，尚不存在一个经过充分测试、验证和公认的机制，用以识别对人工智能文本生成技术的不当应用。因此，在缺乏有效监管的环境中，错误地将由人工智能工具如 ChatGPT 生成的文本当作原创内容提交，仍然较为便利。为了应对这一挑战，开放人工智能公司开发了一种经过训练可以区分人工智能书写的文本和人类书写的文本的免费工具（人工智能文本分辨器）。不幸的是，开放人工智能公司将其描述为“不完美的工具”，并警告称不应将其用作主要决策工具。学术机构和出版商任然不确定是否使用开放人工智能公司不完善的工具或开发更好的工具。

3.2.考虑在学术界负责任地使用 ChatGPT 和其他生成式人工智能系统的可行步骤

在学术领域，负责任地运用生成式人工智能技术至关重要，关键在于如何平衡其优势与对学术诚信可能带来的风险。首要任务应是明确界定：在学术作品中使用人工智能生成的文本而不明确标注来源，应被视为学术不端行为。尽管这在当前的学术诚信政策中有所暗示，但需要让老师、学生以及期刊、潜在作者更明确，当人工智能生成的文本被用作原创作品时，会违反哪些价值观。此外，ChatGPT 和其他人工智能文本生成器可以通过多种方式融合到从评估、研究到教学的学术活动中。学术界对这些方法的了解仍然很少。生成式人工智能系统正在重塑学生所经历的世界格局。因此，学术机构有责任让学生为大语言模型正在有效变革的世界做好准备。

相关机构有必要对教师和学生进行关于 ChatGPT 不同应用场景的培训。对于期望识别学术不诚信工具的使用的老师首先应该能够知道这些工具是如何工作的。

进一步而言，高校若希望维持现有的以书面作业为基础的评价体系，便迫切需要开发一款高效的软件，用以识别由 AI 撰写的内容。然而，设计这样一个工具并将其纳入有效、可靠的评估方法将需要大量资金以及收购开放人工智能公司或这些语言模型的其他创建者或者得到他们的支持。开发可能也需要时间，而与此同时，人工智能生成的文本可能已经成为学术评估的一部分。另一方面，这可能是一个重新考虑论文写作作为一种评估形式的未来的机会。有学者提出，需要对评估体系进行彻底的改革，即不再仅仅关注最终完成的论文成果，而是更加重视在写作过程中所展现的批判性思维能力。教学生成为优秀的写作者很重要，但理解过程不是比完成作品更重要吗？集成类似 ChatGPT 的人工智能工具于学生批判性思维及写作训练中，是一个值得考虑的策略。如果论文是绝对必要的，答辩可以更多地用于更好地进行评估的补充。

这不是一个仅限于学术论文和学生的问题。这种情况在更广泛的学术活动中也有风险：期刊文章、学术会议的论文、研究报告、在线博客文章、学位论文以及书籍等。然而，有一种观点认为，该系统可以让人们发挥自己的优势，提高学术成果的质量。尽管存在种种局限和缺陷，ChatGPT 可以成为一个有效的学习伙伴。例如，它可以产生伟大的想法和文本，这些想法和文本反过来可以被用户完善。它不是权威的学术之声，也不是 100%准确的，但它可以成为一个很好的学术助手。因此，设计采用 ChatGPT 时引用的方法是确保其在学术界负责任使用的必要条件。对于那些希望在开发识别作弊的技术工具之前保持学术诚信原则的用户来说，引用 ChatGPT 可能涉及内容生成日期、内容生成的提示，并将直接引用的使用限制在一段内。

此外，ChatGPT 编写或更正代码的可能性要求重新构想技术编码评估。到目前为止，它已经被证明能够用自定义提示编写功能代码，这可以帮助学生回答基本的数据结构和算法问题。因此，将答辩作为评估的一部分不应仅是补充，而应是评估的主要部分。这将有机会测试学生对代码及其功能的掌握。

4.结论

ChatGPT 和其他人工智能文本生成器的使用方式会破坏学术诚信，也有能力彻底改变学术界。所有人都有责任最大程度地降低它们对学术诚信的风险。这一过程需要众多利益相关者的积极参与和协作；包括但不限于技术研发专家、高等教育机构的决策者、学术出版界、教育工作者如教授和讲师，以及学习者如学生。学术写作、论文作业和技术编码评估可能还

没有消亡，但现在是时候重新构想关键变革，以确保学术界的可持续的诚信了。

总之，学术机构需要做很多事情：

•接受 ChatGPT 作为教育和研究的重要组成部分。

•为老师和学生搭建 ChatGPT 培训，并确保负责任地使用。向教职员工和学生提供必要的支持和资源有助于降低学术诚信可能面临的风险。

•审查学术诚信政策，并做出必要的改变，以反映当前人工智能的趋势和可能性。

•与相关机构（包括但不限于期刊编辑和出版商）合作，共同创造对在学术文本中使用 ChatGPT 和其他人工智能工具做出引用的有效方式。

•致力于开发性价比高且值得信赖的工具，以识别全球学术界可能不诚实地使用人工智能工具的情况。

最后，开放人工智能公司和其他大型语言模型创建者应该愿意与学术界合作，在学术界负责任地使用人工智能驱动的文本生成器。开放人工智能公司开发“不完美”识别器的举措是一个受欢迎的发展，但不足以解决学术诚信问题。该公司目前与美国教育工作者的合作也值得称赞。然而，这种参与应该扩大到世界其他地区学术界的利益攸关方，特别是中低收入国家的利益攸关者。维护学术诚信需要各利益方的协同努力，共同设计出有效的策略来保障学术的正直性。这可能包括重新定义什么是学术成就、影响和衡量它们的新方法。

参考文献

[1] 我的 AI 助手 ChatGPT 日常应用指南[M].刘凯，译.北京：中国科学技术出版社，2024:15.
[2] Choi, J.H., Hickman, K.E., Monahan, A., Schwarcz, D., ChatGPT goes to law School[R]. 2023. Available SSRN.
[3] Terwiesch, C., Would chat GPT3 get a Wharton MBA? A prediction based on its performance in the operations management course[R]. 2023.Mack Inst. Innov. Manag. Whart.
Sch. Univ. Pa.
[4] International Center for Academic Integrity (2021) The fundamental values of academic
integrity, 3rd edn[A]. Available at: https://academicintegrity.org/resources/fundamental-values (访问日期：22 年 10 月 10 日).
[5] Newton, P. M. How common is commercial contract cheating in higher education
and is it increasing? A systematic review. Frontiers in Education, 3. [J] (2018).https://doi.org/
10.3389/feduc.2018.00067
[6] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. On the dangers of stochastic parrots: Can language models be too big?. In Proceedings of the 2021 ACM Conference on fairness, accountability, and transparency, FAccT ’21 (pp. 610 – 623).New York, NY, USA: Association for Computing Machinery. [C] (2021).https://doi.org/10.1145/3442188.3445922.
[7] Stokel-Walker, C. ChatGPT listed as author on rese arch papers: Many scientists disapprove. [J] Nature, 613, 620 – 621. (2023).
https://doi.org/10.1038/d41586-023-00107-z